

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



THOMSON  
DELPHION

RESEARCH  
My Account | Products

PRODUCTS  
Search: QuickNumber Boolean Advanced Devwent

INSIDE DELPHION  
Help

Log On! | Word Files | Saved Searches

The Delphion Integrated View

Buy Now: ☒ PDF | [More choices...](#)

Tools: Add to Work File: ☐ Create new Work File

View: INPADOC | Jump to:

☒ Email this to a friend

Title: JP11353124A2: DISK ARRAY DEVICE

Country: JP Japan  
Kind: A2 Document Laid open to Public inspection !  
Inventor: TAKATSUKI TAKECHIYO;  
Assignee: ALPS ELECTRIC CO LTD  
News, Profiles, Stocks and More about this company

Published / Filed: 1999-12-24 / 1998-06-11

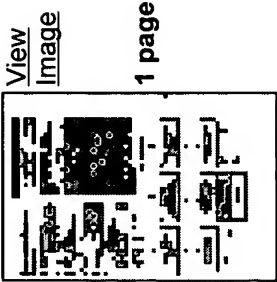
Application Number: JP1998000163637

IPC Code: G06F 3/06; G06F 3/06; G11B 20/12;

Priority Number: 1998-06-11 JP1998000163637

Abstract: PROBLEM TO BE SOLVED: To improve the data processing capability by increasing the generation speed of parity data and the repair speed of faulty data by making the depth of a stripe as a set of strips of respective member storage devices smaller than the one-sector size of the member storage devices.

SOLUTION: In respective member disks A to D, their data storage areas are divided into stripe units and a virtual disk is formed by making the sets of strips corresponding to respective data recording areas of those member disks A to D as the stripe units. The stripe depth of the strips being the sets of stripes of the respective member disks A to D is made larger than the memory bit width of a memory 5 and less than the one-sector size of the respective member disks A to D. This constitution can reduce operation for changing the designation of row addresses when



View  
Image

1 page

parity data are generated and when faulty data are repaired based on the parity data.

COPYRIGHT: (C)1999,JPO

& Family:

None

& Other Abstract  
Info:

DERABS G2000-121241 DERABS G2000-121241



Nominate this for the Gallery...

© 1997-2004 Thomson

Research Subscriptions | Privacy Policy | Terms & Conditions | Site Map | Contact Us | Help

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-353124

(43) 公開日 平成11年(1999)12月24日

(51) Int.Cl.<sup>6</sup>

G 0 6 F 3/06

G 1 1 B 20/12

識別記号

3 0 5

5 4 0

F I

G 0 6 F 3/06

G 1 1 B 20/12

3 0 5 C

5 4 0

審査請求 未請求 請求項の数 5 O L (全 12 頁)

(21) 出願番号

特願平10-163637

(22) 出願日

平成10年(1998)6月11日

(71) 出願人 000010098

アルプス電気株式会社

東京都大田区雪谷大塚町1番7号

(72) 発明者 高月 武千代

東京都大田区雪谷大塚町1番7号 アルプス電気株式会社内

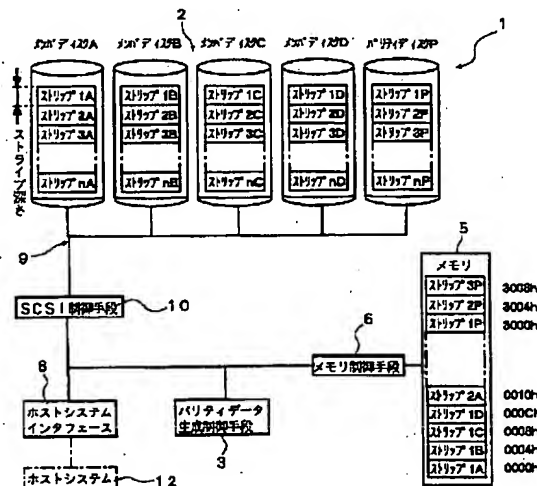
(74) 代理人 弁理士 中尾 俊輔 (外2名)

(54) 【発明の名称】 ディスクアレイ装置

(57) 【要約】

【課題】 パリティデータ生成時や故障データ修復時に、メモリのアドレスマッピングに影響されずにページミスによるオーバーヘッド分の時間ロスを減少し、パリティデータ生成速度や故障データ修復速度を速くしてデータ処理能力を向上させるディスクアレイ装置を提供すること。

【解決手段】 各メンバ記憶装置A, B, C, DのストリップnA, nB, nC, nDの集合であるストライプのストライプ深さをメモリ5のメモリビット幅のサイズ以上であって前記各メンバ記憶装置A, B, C, Dの1セクタサイズより小さくする。



## 【特許請求の範囲】

【請求項1】 ストリップ単位のアドレスに細分化されたデータ記録領域を有する複数のメンバ記憶装置と、これらのメンバ記憶装置のストリップアドレスから読み出すデータおよび前記ストリップアドレスに書き込むデータを一時的に格納するメモリとを有し、前記各メンバ記憶装置に並列にアクセスして処理を行うディスクアレイ装置であって、前記各メンバ記憶装置のストリップの集合であるストライプの深さを前記メモリのメモリビット幅のサイズ以上であって前記メンバ記憶装置の1セクタサイズより小さくしたことを特徴とするディスクアレイ装置。

【請求項2】 前記ストライプ深さを前記メモリビット幅と同一としたことを特徴とする請求項1に記載のディスクアレイ装置。

【請求項3】 前記各メンバディスクへの接続を制御する通信制御手段を有するとともに、前記各メンバディスクと前記メモリとの間のデータ転送を行う場合に前記通信制御手段から出力されるメモリアドレスがどのメモリオフセットアドレスを要求しているかを判断して前記メモリに所定の単位でインクリメントするメモリアドレスに変換するメモリ制御手段を有することを特徴とする請求項1または請求項2に記載のディスクアレイ装置。

【請求項4】 前記メモリ制御手段は、仮想メモリのアドレスのうち上位のメモリアドレスを使用してスキャッタ・ギャザ転送を行うか否かを判断するようにしたことを特徴とする請求項3に記載のディスクアレイ装置。

【請求項5】 前記メモリ制御手段は、前記仮想メモリのアドレスのうち前記各メンバディスクにおけるアドレスの指定に不要なアドレスビットを使用して前記各メンバディスクの番号を指定するようにしたことを特徴とする請求項3または請求項4に記載のディスクアレイ装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明はディスクアレイ装置に係り、特に、メンバ記憶装置が故障して読み取り不能となった故障データを修復するためのパリティデータを生成するのに好適なディスクアレイ装置に関する。

## 【0002】

【従来の技術】 従来から、記憶装置の性能を改善することを目的として、複数のハードディスクドライブ（以下、「メンバディスク」という）を並列のインタフェースを介して接続することによりアクセス時間の短い大容量のディスクサブシステムを構築するディスクアレイ装置が開発されている。

【0003】 このようなディスクアレイ装置によれば、データを各メンバディスクに分散して記録するので並列に高速アクセスすることが可能となるし、パリティディスクを設けておけば、1台のメンバディスクが壊れても

パリティデータにより故障データを修復することも可能となり、故障ディスクを新しいメンバディスクに交換することにより故障ディスクのデータを復旧することも可能である。

【0004】 また、非常用のスペアディスクを装備しておけば、1台のメンバディスクが故障してもスペアディスク上に故障ディスクのデータを生成し、故障ディスクをシステムから切り離すことができる。

【0005】 このような動作を実行するために、図11に示すように、従来のディスクアレイ装置21においては、前記各メンバディスクA、B、Cのデータ記録領域が連続するアドレスを指定した同一サイズのブロックのグループ（以下、「ストリップ」と呼ぶ）に細分化され、これらのストリップ1A、2A、・・・、10C、のうち仮想ディスクを構成する各メンバディスクA、B、C、のデータ記録領域に対応するストリップ1A、1B、1C、の集合がストライプ単位として管理されていた。ここで、前記仮想ディスクとは、アレイ管理ソフトウェア（図示せず）により前記各メンバディスクA、B、Cのストリップ1A、2A、・・・、10C、に基づいて想定される抽象的なディスクのことである。また、前記各ストリップ1A、2A、・・・、10C、内のブロック数は、データ記録領域のストリップサイズと呼ばれ、前記ストライプ1、2、・・・、10のストライプ深さに相当するものである。

【0006】 また、図12に示すように、前記ディスクアレイ装置21には、パリティデータ1P、2P、・・・、nPを格納するパリティディスクPが配設されているとともに、前記仮想ディスクのアドレスに書き込む書込データやその仮想ディスクのアドレスから読み出す読出データ、さらには前記パリティデータ1P、2P、・・・、nPを一時的に格納するためのメモリ22が装備されている。

【0007】 このため、前記ディスクアレイ装置1は、前記各メンバディスクA、B、C、Dに書込データを書き込む場合には、前記メモリ22に一時的に格納してある各ストリップ1A、2A、・・・、nDの書込データを前記各メンバディスクA、B、C、Dに格納するとともに、故障データの修復に備えて、前記メモリ22上の各ストリップデータからパリティ関数（排他的論理和）を使ってパリティデータ1P、2P、・・・、nPを生成し、前記パリティディスクPに格納するようになっている。

【0008】 また、例えば、メンバディスクCのデータ1Cが故障した場合には、前記ディスクアレイ装置21は、前記パリティディスクPからパリティデータ1Pを読み出すとともに他の正常なメンバディスクA、B、Dからデータ1A、1B、1Dを読み出してパリティ関数を使って故障データ1Cを修復するようになっている。

【0009】 ところで、前述した従来のディスクアレイ

装置21においては、前記各メンバディスクA、B、C、Dのストリップサイズが、ハードディスクの1セクター分に相当する512バイトかあるいは512バイトの整数倍の大きさに細分化されていた。例えば、前記ストリップサイズが512バイトであるとする、このストリップ1A、1B、1C、1D、・・・、nDに対して書き込むデータあるいは読み出すデータが前記メモリ22に格納される場合には、図13に示すように、C00h、E00h、1000h、・・・と512バイト単位ごとにアドレスが増加されていた。そして、前記メモリに格納された各書込データをパリティ関数(XOR)に代入してパリティデータ1P、2P・・・nPを生成し、これを前記メモリ22の所定のアドレスに格納するようになっていた。

#### 【0010】

【発明が解決しようとする課題】しかし、従来のディスクアレイ装置21においては、前記各メンバディスクA、B、C、Dのストリップサイズが512バイトあるいは512×(整数)バイトの大きさに形成されていたため、パリティデータ1P、2P、・・・、nPを生成したときに、メモリのページミスにより極めて大きなオーバーヘッド分の時間ロスが生じる場合があった。

【0011】すなわち、前記ストリップサイズ(ストリップデータのサイズ)が512バイトであって、前記メモリが、10ビットのロウアドレスと10ビットのカラムアドレスを有するダイナミックRAM(DRAM)のファーストページモードあるいはEDOモードとし、そのメモリビット幅が32ビット(4バイト)で構成されているとすると、パリティデータ1P、2P、・・・、nPを生成するときのメモリアクセスは図14に示すようなタイミングチャートとなっていた。ここで、図14中のRAS信号およびCAS信号は、それぞれロウアドレスストロブおよびカラムアドレスストロブのことであり、これら2種類の信号線により前記DRAMのアドレスデータを指定するようになっている。

【0012】そして、図13に示すように、従来のディスクアレイ装置21では、ストリップA1(C00h~DFFh)からストリップB1(E00h~FFFh)への移行時、あるいはストリップC1(1000h~1FFFh)からストリップD1(1200h~13FFh)への移行時には、ロウアドレスがそれぞれ「0」から「0」、あるいは「1」から「1」となり変更されないため、最初にロウアドレスにおいて上位アドレスを「0」あるいは「1」とラッチしておけば、逐一ロウアドレスを入力する必要はない。

【0013】しかし、ストリップB1(E00h~FFFh)からストリップC1(1000h~1FFFh)へ移行するときには、ロウアドレスが「0」から「1」へと変更されてロウアドレスデータを改めてラッチしなければならなかった。このため、一旦、RAS信号をH

ighレベル(Hi)にしてロウアドレス「1」を入力し、その後、HiからLowレベル(Lo)に戻す必要があった。

【0014】したがって、このRAS信号の操作が、ページミスによるオーバーヘッドとなり、パリティデータ1P、2P・・・、nPを生成する際の時間ロスを生じてしまっていた。この時間ロスは、例えば、1回のオーバーヘッド時間が100ナノ秒(ns)であれば、ストリップP1のパリティデータを生成するために、512/4(ストリップデータサイズ/メモリビット幅)=128回のパリティデータ生成動作を行わなければならない、全体として100ns×128=12.8マイクロ秒(μs)となる。

【0015】つまり、従来のディスクアレイ装置21においては、ストリップ1A、2A・・・、nDの書込データあるいは読出データが、ロウアドレスデータを変更するアドレスに格納された場合には、1ストリップ当たりのパリティデータnPを生成するために12.8μsの時間ロスが生じてしまうこととなっていた。

【0016】このような時間ロスは、前記ストリップデータサイズを1Kバイト以上にさらに大きくした場合に、前述の例にあてはめると、パリティデータnPを作成することに必ずページミスが発生する事態となってしまう。

【0017】本発明はこのような問題点を鑑みてなされたもので、パリティデータを生成する場合や故障データをパリティデータに基づいて修復する場合に、メモリのアドレスマッピングに影響されずにページミスによるオーバーヘッド分の時間ロスを減少でき、パリティデータの生成速度および故障データの修復速度を速くしてデータ処理能力を向上させることのできるディスクアレイ装置を提供することを目的とするものである。

#### 【0018】

【課題を解決するための手段】前記目的を達成するため本発明に係る請求項1に記載のディスクアレイ装置の特徴は、各メンバ記憶装置のストリップの集合であるストライプのストライプ深さをメモリのメモリビット幅のサイズ以上であって前記各メンバ記憶装置の1セクタサイズより小さくした点にある。そして、このような構成を採用したことにより、パリティデータを生成する場合や故障データをパリティデータに基づいて修復する場合に、ロウアドレスの指定を変更する動作を減少できるため、前記メモリのアドレスマッピングの仕方に影響されずにページミスによるオーバーヘッド分の時間ロスを減少し、パリティデータの生成速度および故障データの修復速度を速くしてデータ処理能力を向上させることができる。

【0019】また、請求項2に記載のディスクアレイ装置の特徴は、請求項1において、ストライプ深さをメモリビット幅と同一とした点にある。そして、このような

構成を採用したことにより、パリティデータを生成する場合や故障データをパリティデータに基づいて修復する場合に、ページミスによるオーバーヘッド分の時間ロスを最大限に減少させることができ、より一層確実にパリティデータの生成速度および故障データの修復速度を速くすることができる。

【0020】また、請求項3に記載のディスクアレイ装置の特徴は、請求項1または請求項2において、各メンバディスクへの接続を制御する通信制御手段を有するとともに、前記各メンバディスクとメモリとの間のデータ転送を行う場合に前記通信制御手段から出力されるメモリアドレスがどのメモリオフセットアドレスを要求しているかを判断して前記メモリに所定の単位でインクリメントするメモリアドレスに変換するメモリ制御手段を有する点にある。そして、このような構成を採用したことにより、メンバディスクとメモリとの間のデータ転送を円滑かつ確実に行うことができるため、パリティデータの生成やパリティデータに基づく故障データの修復を効果的に実行できる。

【0021】また、請求項4に記載のディスクアレイ装置の特徴は、請求項3において、メモリ制御手段が、仮想メモリのアドレスのうち上位のメモリアドレスを使用してスキップ・ギャザ転送を行うか否かを判断するようにした点にある。そして、このような構成を採用したことにより、連続するアドレスをアクセスするデータ転送か不連続のアドレスをアクセスするデータ転送かを明確に判断できて、メンバディスクとメモリとの間、若しくはホストシステムとメモリとの間のデータ転送を円滑かつ確実に行うことができる。

【0022】また、請求項5に記載のディスクアレイ装置の特徴は、請求項3または請求項4において、メモリ制御手段が、仮想メモリのアドレスのうち各メンバディスクにおけるアドレスの指定に不要なアドレスビットを使用して前記各メンバディスクの番号を指定するようにした点にある。そして、このような構成を採用したことにより、各メンバディスクの指定を効率的かつ確実に指定することができて、メンバディスクとメモリとの間のデータ転送を円滑かつ確実に行うことができる。

【0023】

【発明の実施の形態】以下、本発明の実施の形態を図1乃至図10を参照して説明する。

【0024】本実施形態のディスクアレイ装置1は、図1に示すように、複数のメンバ記憶装置の例としてメンバディスクA、B、C、Dから構成されるメンバディスク群2と、これらのメンバディスクA、B、C、Dの故障データを修復するためのパリティデータが格納されるパリティディスクPと、前記各メンバディスクA、B、C、Dに書き込まれる書込データに基づいてパリティデータを生成し前記パリティディスクに格納する制御を行うパリティデータ生成制御手段3と、前記メンバディス

クの故障データを前記パリティデータに基づいて修復する制御を行う故障データ修復制御手段4と、ホストシステム12からのアクセスにより前記各メンバディスクA、B、C、Dに書き込まれる書込データおよび前記各メンバディスクA、B、C、Dのアドレスから読み出される読出データならびに前記パリティデータを一時的に保管するメモリ5と、このメモリ5へ前記各データを保管するための制御を行うメモリ制御手段6と、ホストシステム12およびディスクアレイ装置1間の仲介を行うホストシステムインタフェース8と、前記メンバディスク群2へのアクセス伝達経路であるSCSIバス9の制御を行うSCSI制御手段10とから構成されている。

【0025】以下、これらの各手段についてより具体的に説明する。

【0026】前記メンバディスク群2の各メンバディスクA、B、C、Dは、不揮発でランダムにアドレスが設定できて書き換え可能な大容量記憶装置であって、例えば、磁気ディスク、光ディスクおよび半導体ディスク等が相当する。

【0027】なお、本実施形態における前記メンバディスク群2は、4台のメンバディスクA、B、C、Dにより構成されているが、この台数に限る必要はなく、用途に応じて増減させてよい。

【0028】また、前記各メンバディスクA、B、C、Dは、図1に示すように、そのデータ記録領域がストリップ単位に細分化されており、これらのメンバディスクの各データ記録領域に対応するストリップの集合をストライプ単位として仮想ディスクが構成されている。例えば、仮想ディスクのストライプ1は、ストリップ1A、ストリップ1B、ストリップ1Cおよびストリップ1Dの集合により構成されている。そして、このストライプの深さは、前記メモリ5のメモリビット幅と同一であるか、あるいは、同一でなくともパリティデータ生成の際にメモリマッピングに影響されずにオーバーヘッド分の時間ロスを減少させられる大きさとされている。すなわち、前記ストライプ深さがメモリビット幅より大きくてもそのメモリビットの非常に小さい整数倍であれば、オーバーヘッド分の時間ロスを十分減少できるようになっている。

【0029】ただし、前記ストライプ深さがメモリビット幅よりも大きくなるに従ってページミスの発生率が高くなるため、このストライプ深さは、メンバディスクの台数や前記メモリ5のDRAMの構成等に応じて適正に定める必要がある。

【0030】また、前記ストライプ深さに関してその他の具体例を挙げて説明すると、前記メンバディスクが4台、前記メモリ5が10ビットのロウアドレスと10ビットのカラムアドレスを有するDRAMで32ビット幅とされて場合に、ストライプ深さがその1Kより大きい場合には、常に、ページミスが発生するため、このスト



ライプ深さの大きさは適当ではない。

【0031】また、ストライプ深さが従来のように512バイトとした場合であっても、例えば、メモリマッピングがストリップ1Aを000h~1FFhとし、ストリップ1Bを200h~3FFhとし、ストリップ1Cを400h~5FFhとし、ストリップ1Dを600h~7FFhとするなどのように、メモリマッピングのやり方によってはページミスが発生させずにすむ。しかし、ページミスが発生するメモリマッピングにした場合に、従来技術において説明したように、512バイトのパリティデータを生成するための各データの読み出しループごとにページミスが発生してしまい、膨大に時間ロスを生じてしまうため妥当でない。

【0032】これに対して、前記ストライプ深さをメモリビット幅である4バイトにすると、図2に示すように、ページミスが発生するのはストリップ64Bからストリップ64Cの位置だけであり、ページミス発生率は極めて少なく、それによる時間ロスの影響はない。

【0033】また、前記パリティディスクは、前記各メンバディスクの各ストリップに書き込まれる書込データ（「ストリップデータ」ともいう）に基づいて前記パリティ生成制御手段により生成されるパリティデータを格納するようになっている。

【0034】このため、このパリティディスクは、前記メンバディスクと同様に、ストリップ単位に細分化されている。

【0035】前記パリティデータ生成制御手段3は、前記各メンバディスクA、B、C、Dのストリップデータをそれぞれ非論理和のパリティ関数に代入し、パリティデータを生成するようになっている。そして、このように生成したパリティデータを前記パリティディスクのストリップに順次格納するようになっている。

【0036】前記故障ディスク修復制御手段は、前記メンバディスクのいずれか1つのディスクが故障したために、損失したデータを前記パリティデータおよび他の正常なメンバディスクのストリップデータをパリティ関数に代入して修復するようになっている。

【0037】前記メモリ5は、ホストシステム12がメンバディスクとアクセスしたときに、ストリップデータおよびパリティデータを一時的に保存するようになっている。例えば、ホストシステム12からメンバディスクに書込データが送信されてくると、当該メモリ5に一時的に保存し、前記パリティデータ生成制御手段3がこの書込データに基づいてパリティデータを生成するようになっている。あるいは、メンバディスクのストリップデータが故障した場合に、他の正常なメンバディスクから読み出されるストリップデータおよび前記パリティディスクから読み出されるパリティデータを当該メモリ5に一時的に保存し、前記故障データ修復制御手段4がこれらの各データに基づいて故障データを修復するようにな

っている。

【0038】また、前記メモリ5は、記憶領域を所定のメモリビット幅のサイズに細分化されており、このメモリビット幅にあわせて前記各メンバディスクのストリップデータおよびパリティデータが一時的に保存されるようになっている。

【0039】前記メモリ制御手段6は、前記SCSI制御手段10を介して前記各メンバディスクと前記メモリ5との間のデータ転送を行う場合に、前記SCSI制御手段10から出力されるメモリアドレスがどのメモリオフセットアドレスを要求しているかを判断し、所定の単位で前記メモリ5にインクリメントするメモリアドレスに変換するようになっている。すなわち、例えば、図1に示すように、前記メンバディスクのストライプ深さが4バイトである場合において、メンバディスクAと前記メモリ5との間のデータ転送は、オフセットアドレス0000hに対して0010h、0020h、・・・と10h単位のメモリアドレスにアクセスし、メンバディスクBと前記メモリ5との間のデータ転送は、オフセットアドレス0004hに対して0014h、0024h、・・・と10h単位のメモリアドレスにアクセスするようになっている。一方、SCSI制御手段10から出力されるメモリアドレスは、1回のデータ転送単位ごと、本実施形態においては4h単位ごとにインクリメントされる。

【0040】これらのデータの転送およびアドレスの変換には、ストリップナンバー信号およびスキップ・ギャザ転送信号が必要となる。ここで、ストリップナンバー信号とは、前記各メンバディスクが特定のストリップデータを転送させるために前記メモリ5のどのオフセットアドレスを要求するかを示す信号である。また、スキップ・ギャザ転送信号とは、前記メンバディスク上で隣接して記録されているストリップデータを隣り合わないメモリアドレスに転送すること（スキップリード）を示す信号および連続していないメモリアドレスに記録されたストリップデータのブロックを連続したアドレスディスクブロック転送すること（ギャザライト）を示す信号のことをいう。

【0041】前記データの転送およびアドレス変換を行うため、前記メモリ制御手段6は、図3に示すように、前記SCSI制御手段10により要求されたアドレスを受けてスキップ・ギャザインエーブルであるか、ストリップナンバーが何番かをそれぞれ判定するスキップ・ギャザ転送判定ロジック13と、このスキップ・ギャザ転送判定ロジック13から発信されるスキップ・ギャザインエーブル信号およびストリップナンバー信号に基づいてスキップ・ギャザアドレスを生成するスキップ・ギャザアドレス生成回路14と、このスキップ・ギャザアドレスをメモリアドレスに変換して前記メモリ5に転送するとともにRAS信号およびCAS信号を発

信制御するアドレス・制御信号コントローラ15とにより構成されている。

【0042】したがって、前記メモリ制御手段6は、図4に示すように、前記メモリ5が64KBのメモリサイズを有しているとする、実メモリ空間64KBに対して仮想メモリ空間128KBを想定し、上位のメモリアドレスであるアドレスビット16を使用してスキャットギャザ転送をするか否かを判断するようになっている。

【0043】すなわち、アドレスビット16が「0」の場合には、仮想メモリアドレスと実メモリアドレスは、図3中のストリップ1P(03000h=3000h)およびストリップ2P(03004h=3004h)に示すように、同一アドレスとなるためスキャットギャザ転送をしない。一方、アドレスビット16が「1」の場合には、図4中のストリップ1A(10000hと0000h)やストリップ2A(10004hと0010h)等に示すように、同一アドレスではないためスキャットギャザ転送をするようになっている。

【0044】また、前記各メンバディスクと前記メモリ空間とがアクセスする場合、1つのメンバディスクからメモリアクセスできるメモリ空間のメモリサイズは、メンバディスクの台数が4台であれば、 $64\text{KB}/4=16\text{KB}$ となる。このため、各メンバディスクにとって、メモリアドレスの上位2ビットであるメモリビット14、15は、実質的に不要となる。すなわち、メンバディスクAは、仮想メモリ上10000h~13FFFhの範囲に配置されているが、実際にメンバディスクAをハード的にアクセスするときは0000h~3FFFhの範囲であって、アドレスは下位14ビットあればよく、アドレスビット14、15は不要となる。

【0045】そこで、このアドレス指定に不要となった2ビットを使用して、各メンバディスクに対応するストリップナンバーを指定するようになっている。すなわち、ビットが「00」のときはストリップA(メンバディスクA)、「01」のときはストリップB(メンバディスクB)、「10」のときはストリップC(メンバディスクC)および「11」のときはストリップD(メンバディスクD)を指定するようになっており、これにより仮想メモリアドレスは実メモリアドレスに変換されるようになっている。

【0046】つぎに、前記ホストシステムインタフェース8は、ホストシステム12とディスクアレイ装置1との間におけるデータ転送を調整するようになっており、ソフト的にはホストシステム12からのコマンド発行、完了等を制御するもので、前記SCSI制御手段10は、各メンバディスクとメモリとの間におけるデータ転送を調整するもので、ソフト的にはSCSIバス9の制御を行い、各メンバディスクとインタフェースするようになっている。

【0047】つぎに、本発明の実施形態の作用について

説明する。

【0048】本実施形態のディスクアレイ装置1は、ホストシステム12から各メンバディスクに書込データを書き込む命令が出されると、前記ホストインタフェースがそのコマンドを制御して書込データを受けとる。この書込データは、前記メモリ制御手段6を経由してメモリ5に保存される。また、この書込データは、ストライプ深さに分割されて各メンバディスクに割り振られることとなるが、その前に前記パリティデータ制御手段が、前記各ストリップデータをパリティ関数に代入することによりパリティデータを生成する。このパリティデータは前記メモリ制御手段6の制御により一時的に前記メモリ5に保存されることとなる。その後、前記SCSI制御手段10が、前記ストリップデータおよび前記パリティデータを各メンバディスクのストリップアドレスに格納するとともに、前記パリティデータをパリティディスクに格納する。

【0049】一方、前記メンバディスク群2の中の1つのディスクが故障し、そのストリップデータが損失した場合には、本実施形態のディスクアレイ装置1は、前記SCSI制御手段10を起動させて、故障ディスク以外の正常なディスクのストリップデータと前記パリティデータを読み出す。これらのストリップデータおよびパリティデータを前記メモリ制御手段6がメモリアドレスに変換して前記メモリ5に一時的に保存する。そして、パリティ生成コントローラ3内のパリティデータ制御手段が、前記ストリップデータおよび前記パリティデータを前記パリティ関数に代入することにより故障したストリップデータを修復する。この修復されたデータは、一旦前記メモリ5に保存された後に、前記メモリ制御手段6から前記ホストI/F経由でホストシステムに転送され、正常データとして扱われる。

【0050】つぎに、図5乃至図10において、本実施形態のディスクアレイ装置1の前記メモリ制御手段6が、前記各メンバディスクのストリップデータを前記メモリ5に転送して得られるメモリアドレスマッピングおよびこれらのストリップデータに基づいて前記パリティデータ生成制御手段3がパリティデータを生成する際のメモリアクセスのタイミングチャートを具体例を示して説明する。

【0051】図5および図6に示すように、前記メモリ5のメモリビット幅が32ビットであり、前記メモリ制御手段6により想定される仮想ディスクのストライプ深さが、そのメモリビット幅32の大きさと同一に設定されている場合には、0h、4h、8h、・・・と4バイトごとにメモリアドレスが増加される。また、パリティデータを生成する際に、マッピングの仕方に影響されずにページミスの発生を減少でき、通常ロウアドレスを再指定する必要はなくオーバーヘッド分の時間ロスを極めて効果的に抑制できる。

【0052】また、図7および図8に示すように、前記メモリ5のメモリビット幅が64ビットであり、前記仮想ディスクのストライプ深さが、そのメモリビット幅64ビットの大きさと同一に設定されている場合には、0h、8h、10h、・・・と8バイトごとにメモリアドレスが増加される。また、パリティデータを生成する際には、マッピングの仕方に影響されずにページミスの発生が減少でき、ロウアドレスを再指定する必要はなくRAS信号を逐一LoからHiに切換えずに済むため、オーバーヘッド分の時間ロスを効果的に抑制できる。

【0053】また、図9および図10に示すように、前記メモリ5のメモリビット幅が32ビットであり、前記仮想ディスクのストライプ深さが、DRAMのページミスの発生が少ない大きさに設定されている場合には、0h、4h、8h、・・・と4バイトごとにメモリアドレスが増加され、メンバディスクAのストリップ1A-0からストリップ1A-3までの16バイト分のアドレスマッピングが行われると、次のメンバディスクBのストリップデータ(1B-0～1B-3)のアドレスマッピングに移行する。また、パリティデータを生成する際には、パリティデータ1P-0からパリティデータ1P-3の生成にかけて通常マッピングの仕方に影響されずにページミスの発生が減少でき、RAS信号の切換えを必要としないため、オーバーヘッド分の時間ロスを効果的に抑制できる。

【0054】したがって、本発明の実施形態によれば、パリティデータを生成する場合や故障データをパリティデータに基づいて修復する場合に、メモリ5のアドレスマッピングの仕方に影響されずにページミスによるオーバーヘッド分の時間ロスを減少し、パリティデータの生成速度および故障データの修復速度を速くすることができる。このため、ディスクアレイ装置1のデータ処理能力を向上させることができる。

【0055】なお、本発明は前記実施の形態のものに限定されるものではなく、必要に応じて種々変更することが可能である。

【0056】

【発明の効果】以上述べたように本発明に係る請求項1に記載のディスクアレイ装置によれば、パリティデータを生成する場合や故障データをパリティデータに基づいて修復する場合に、ロウアドレスの指定を変更する動作を減少できるため、メモリのアドレスマッピングの仕方に影響されずにページミスによるオーバーヘッド分の時間ロスを減少し、パリティデータの生成速度および故障データの修復速度を速くしてデータ処理能力を向上させることができる。

【0057】また、請求項2に記載のディスクアレイ装置によれば、請求項1に記載の発明の効果に加えて、パリティデータを生成する場合や故障データをパリティデータに基づいて修復する場合に、ページミスによるオー

バーヘッド分の時間ロスを最大限に減少させることができ、より一層確実にパリティデータの生成速度および故障データの修復速度を速くすることができる。

【0058】また、請求項3に記載のディスクアレイ装置によれば、請求項1または請求項2に記載の発明の効果に加えて、メンバディスクとメモリとの間のデータ転送を円滑かつ確実に行うことができるため、パリティデータの生成やパリティデータに基づく故障データの修復を効果的に実行できる。

【0059】また、請求項4に記載のディスクアレイ装置によれば、請求項3に記載の発明の効果に加えて、連続するアドレスをアクセスするデータ転送か不連続のアドレスをアクセスするデータ転送かを明確に判断でき、メンバディスクとメモリとの間のデータ転送およびホストシステムとメモリとの間のデータ転送をより円滑かつ確実に行うことができる。

【0060】また、請求項5に記載のディスクアレイ装置によれば、請求項3または請求項4に記載の発明の効果に加えて、各メンバディスクの指定を効率的かつ確実に指定することができ、メンバディスクとメモリとの間のデータ転送を円滑かつ確実に行うことができる。

【図面の簡単な説明】

【図1】 本発明に係るディスクアレイ装置の実施形態を示すブロック図

【図2】 本実施形態における各メンバディスクのストライプ深さをメモリビット幅と同一にしたときにページミスが発生する場合についての説明図

【図3】 本実施形態におけるメモリ制御手段の構成を示す説明図

【図4】 本実施形態におけるメモリ制御手段による実メモリと仮想メモリとの間のデータ転送に関する説明図

【図5】 本実施形態における各メンバディスクのストライプ深さを32ビットのメモリビット幅と同一にした場合のメモリマッピングを示す説明図

【図6】 図5の場合においてパリティデータを生成するときのタイミングチャートを示す説明図

【図7】 本実施形態における各メンバディスクのストライプ深さを64ビットのメモリビット幅と同一にした場合のメモリマッピングを示す説明図

【図8】 図7の場合においてパリティデータを生成するときのタイミングチャートを示す説明図

【図9】 本実施形態における各メンバディスクのストライプ深さを32ビットのメモリビット幅よりも大きい、DRAMリフレッシュより十分小さくした場合のメモリマッピングを示す説明図

【図10】 図9の場合においてパリティデータを生成するときのタイミングチャートを示す説明図

【図11】 従来の一般的なディスクアレイ装置におけるメンバディスクのデータ記録領域の細分化に関する説明図

【図12】 従来の一般的なディスクアレイ装置におけるメンバディスクとメモリとの関係を示す説明図

【図13】 従来のディスクアレイ装置によるメモリマッピングを示す説明図

【図14】 図12の場合においてパリティデータを生  
成するときのタイミングチャートを示す説明図

【符号の説明】

1 ディスクアレイ装置

2 メンバディスク群

3 パリティデータ生成制御手段

4 故障データ修復制御手段

5 メモリ

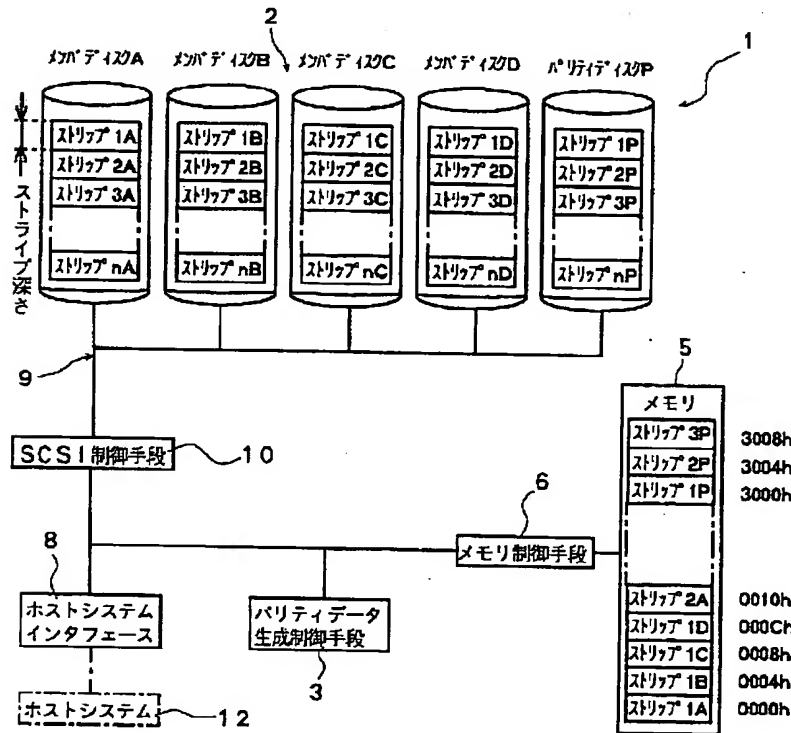
6 メモリ制御手段

13 スキャット・ギャザ転送判定ロジック

14 スキャット・ギャザアドレス生成回路

15 アドレス・制御信号コントローラ

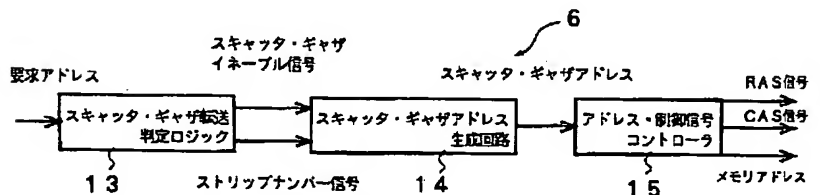
【図1】



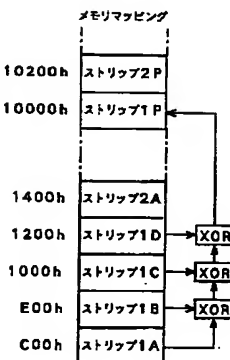
【図2】

1000Ch	ストリップ2P
1008h	ストリップ1P
1404h	ストリップ128D
1400h	ストリップ128C
13FC h	ストリップ128B
13F8h	ストリップ128A
1004h	ストリップ64D
1000h	ストリップ64C
0FFCh	ストリップ64B
0FF8h	ストリップ64A
0C14h	ストリップ1D
0C10h	ストリップ1C
0C0Ch	ストリップ1B
0C08h	ストリップ1A

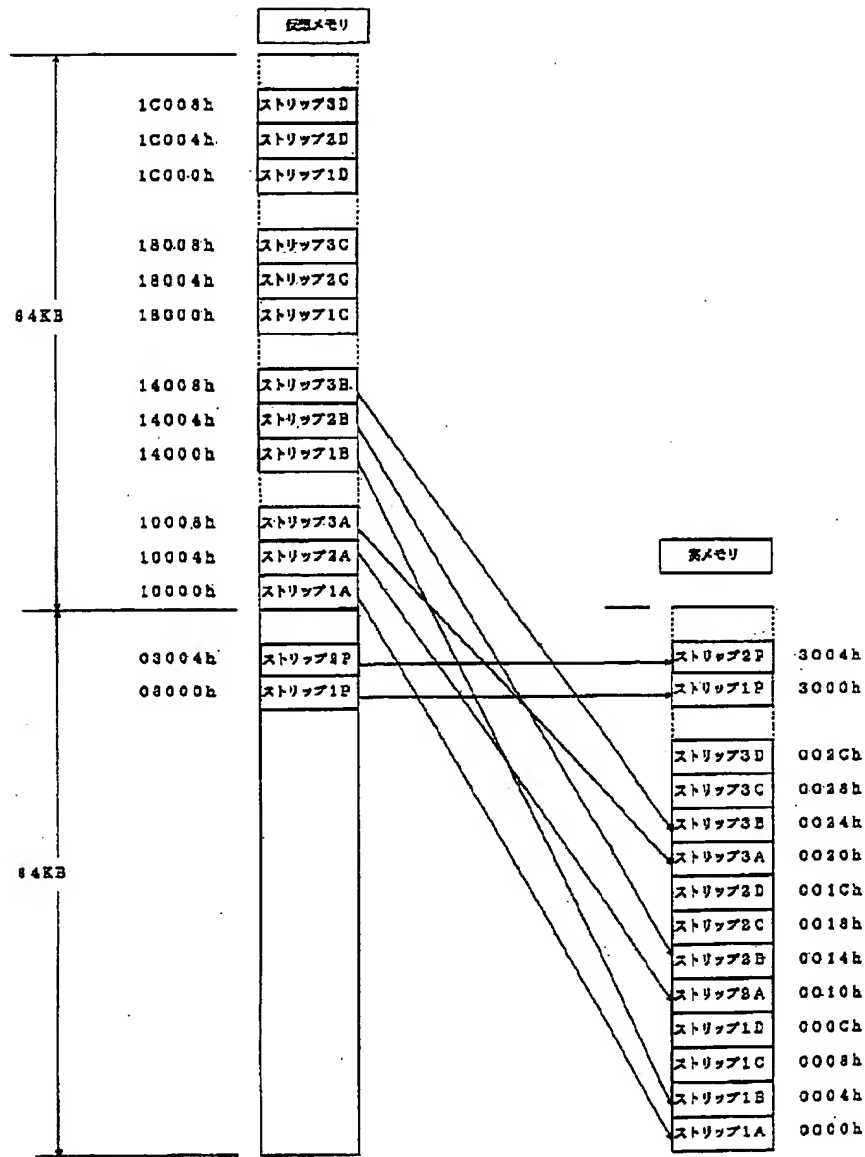
【図3】



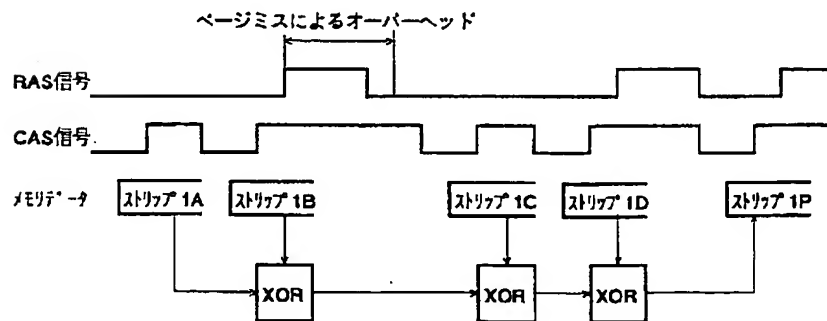
【図13】



【図4】



【図14】



【図5】

メモリビット幅 32ビットの場合

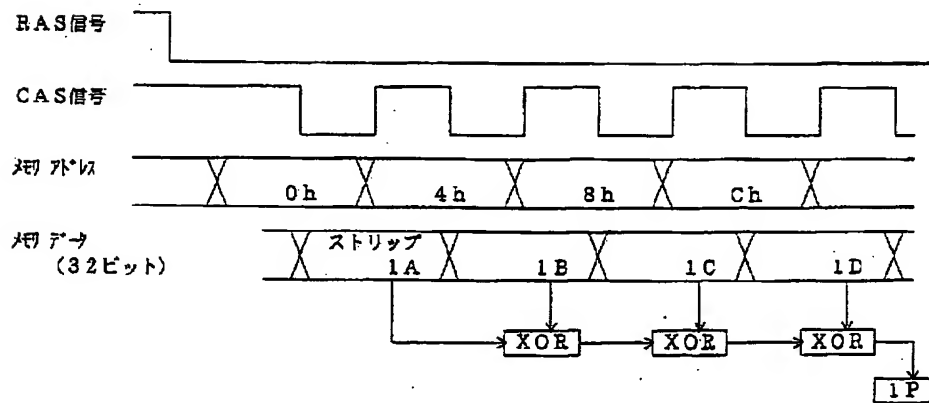
2P	1004h
1P	1000h
...	
3D	30h
3C	28h
3B	24h
3A	20h
2D	1Ch
2C	18h
2B	14h
2A	10h
1D	Ch
1C	8h
1B	4h
ストリップ 1A	0h

【図7】

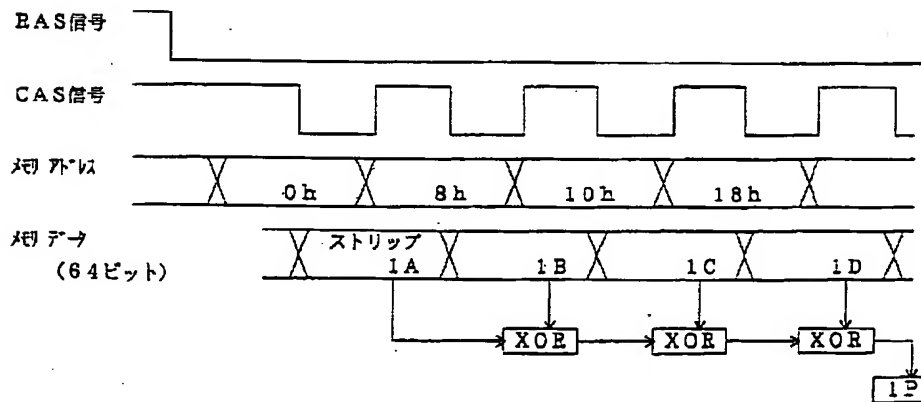
メモリビット幅 64ビットの場合

2P	1008h
1P	1000h
...	
3D	68h
3C	50h
3B	48h
3A	40h
2D	38h
2C	30h
2B	28h
2A	20h
1D	18h
1C	10h
1B	8h
ストリップ 1A	0h

【図6】

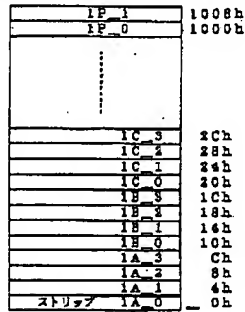


【図8】

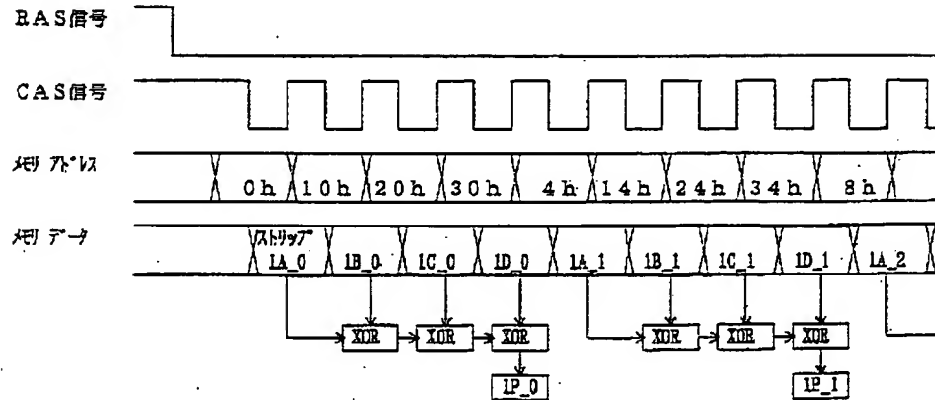


【図9】

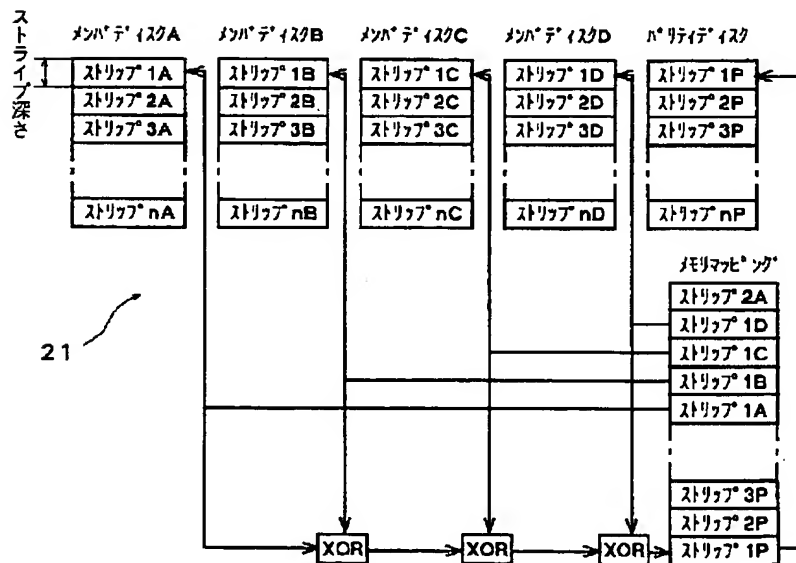
メモリビット幅 32ビットで、かつストライプ幅さが  
ページミスが発生しにくいストライプ幅さの場合



【図10】



【図12】



【図11】

